L. Argabright. Open Access Patterns in Emerging Scholars' Journal Choice: A Bibliometric Analysis. A Master's Paper for the M.S. in IS degree. May, 2020. 59 pages. Advisor: Lorin Bruckner

This bibliometric analysis of University of North Carolina at Chapel Hill research provides a case study for tenure-track faculty's open access publishing action. Looking specifically at the publications of 300 tenure-track assistant professors and instructors, this study creates a series of quantitative information visualizations to see researcher publishing trends leading up to tenure review. It identifies the proportion of assistant professors publishing open access, breaks down the top departments, and shows how open access has increased over time. The study can inform efforts to provide services for tenure-track faculty and emerging scholars. Data from this study was collected through Scopus. Code and interactive visualizations are provided as supplemental files.

Headings:

Bibliometrics

Open access publishing

Scholarly communication

Information visualization

Quantitative research: Content analysis

OPEN ACCESS PATTERNS IN EMERGING SCHOLARS' JOURNAL CHOICE: A BIBLIOMETRIC ANALYSIS

by
L. Argabright

A Master's paper submitted to the faculty
of the School of Information and Library Science
of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements
for the degree of Master of Science in
Information Science.

Chapel Hill, North Carolina

May 2020

Approved by

_____

Lorin Bruckner

Table of Contents

ACKNOWLEDGEMENTS

My advisor, Lorin Bruckner, the Data Visualization librarian at UNC, provided me with a great deal of help throughout the paper. I am sincerely grateful. Her support has helped me not only to scope out my research study and define methods that could appropriately analyze the data, but also to build my skills. Throughout this past year, we have discussed strategies for research reproducibility, information visualization techniques, and ways that programming scripts can quantitatively analyze data instead of manually formatting in Excel. With her guidance, I dramatically improved my Tableau and R programming language capabilities. In each of our repeated encounters, Lorin met me with determination, thoughtfulness, and cheer, which set up a lovely environment for me to learn and be curious.

INTRODUCTION

The open access movement has gained traction since the 1990s as a form of scholarly dissemination. Open access acts as a "free access" option of communicating research findings content for readers: as initially identified in an official capacity by the Budapest Open Access Initiative in 2001, open access is the "free availability [of scholarly literature] on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of the articles, crawl them for indexing, pass them as data to software . . . [without barriers] other than those inseparable from gaining access to the internet itself" (BOAI para 3, as cited in Peekhaus & Proferes, 2015, p. 641). Despite its capacity as "free access," it does not necessarily equate to total "free use," for the author maintains copyrights to the work and thus can loosely control how the work is repurposed and acknowledged in others' work. Conversely, non-open access models of dissemination put restrictions on both access and use—readers must pay to access the material, as well as requesting use of it in their own work.

Open access gives readers and researchers who could not otherwise afford a subscription an opportunity to interact with research. It supports within- and between-community equality, as well as providing a bridge within the global North-South research divide (Tennant, 2016). Libraries are able to provide access to articles despite their declining budgets and correspondingly reduced subscription bundles (Suber, 2012; McKiernan, 2017). Being available to read online without payment restrictions, it

naturally has the opportunity to provide more visibility to the research; with more visibility, can come increased awareness of the author as well as increased use of the work. Open access articles have on average 18% more citations than subscription articles (Piwowar, 2018), which influences an author's impact rating, let alone all the positive societal impacts of this benefit.

Despite these opportunities, open access has grown slowly since its inception. Only about 15% of all scholarly published research exists as open access (University of California, 2019). Numerous possible reasons can be attributed to this slow growth. In part, the open access journals' perceived quality has negatively been affected by newness in age and predatory companies, making researchers reluctant to trust this unfamiliar journal type. Additionally, disciplinary differences cause faculty to engage with publishing models differently. The disciplines vary sharply in their funding availability, which affects researchers' ability to afford a "Gold Open Access" publication, in which the author pays upwards of $5,000 in "Article Processing Charges" (APCs); if an author cannot get a grant, then they would have difficulty finding alternate APC payment sources. Also, university administrators—some fields more than others—are resistant to changing tenure policies that cling to traditionally non-open access publication formats, which can greatly affect a junior researcher's dissemination decisions. These are significant, currently unsolved, areas of discontent about open access.

Junior faculty must fulfill tenure policy requirements within a limited window of time, in preparation for tenure consideration. Due to the nature of one person's permanent tenure security meaning a loss of flexibility for the school to hire an alternate professor, tenure is not awarded merely due to accomplishment; rather, it "should only be made

when there is reasonable confidence that a faculty candidate is one of the best obtainable individuals" (Kenan-Flagler Business School, 2017, p. 3). Tenure-track faculty rank policies at the University of North Carolina at Chapel Hill, for example, allow Instructors 1 year probationary appointments up to 4 times until promotion to Assistant Professor, and Assistant Professors are allowed one 4 year probationary appointment, and one 3 year reappointment if necessary, until promotion to Associate Professor, which would automatically confer tenure.[1] Requirements as to criteria for evaluation of tenure and promotion vary by college and institution, as denoted in the associated group's Faculty Handbook.[2] As these emerging scholars strive to make beneficial dissemination choices for themselves within the publishing world, they may lean on the support and advice of established researchers to increase their own impact, productivity, and visibility. They often develop patterns of behavior in their push to achieve their career goals, regardless of what may be best for the research and research community.

This study aims to examine junior researchers' publication actions, prior to receiving tenure. It briefly identifies collaboration patterns of junior researchers in regards to choice of their open access or non-open access publication format, and strives to connect how publication patterns reveal influences on junior researchers' choices. With a new generation of scholars comes an opportunity to embrace new publication structures. A majority of these emerging scholars have assimilated to an association that information is everywhere, and that research and communication is, for the most part, available and digital. In observance of publication record data, and in connection to various qualitative reports from recent studies reviewed in the Literature Review section,

this study provides patterns of variance between action and beliefs in the junior scholars' path to achieve tenure.

My Research Questions are thus as follows:

1. What is the proportion of Assistant Professors at UNC who are publishing in open access formats?
2. What is this proportion of Assistant Professors at UNC, split up by discipline?
3. How are Assistant Professors at UNC collaborating, in open access versus non-open access publications?

In research question 2, representative data collection and careful data comparison will be necessary. Disciplines have different knowledge production practices, "influenced by different norms, objects, and techniques of study" which can affect productivity and collaboration (Sugimoto et al, 2016, p. 1000). What works for one field could be entirely irrelevant for another, even within the Humanities, Social Sciences, and STEM segments, and thus are not necessarily comparable. Many research studies do lump findings into these Humanities, Social Sciences, and STEM categories.

For the sake of scope, this paper will not elaborate extensively on funding or impact-factor deterrents that impact faculty's choice to publish open access or non-open access articles. These can act as important deterrents to open access dissemination, but each is such a vast topic alone with too much dedicated research regarding it. Useful introductions to each topic can be found in Beasley (2016) and Agarwal et al. (2016) respectively. The next section reviews literature related to rank, tenure preparation, collaboration, and disciplinarity, in an attempt to offer a well-rounded but concise view

of the open access landscape. It will then be followed with a descriptive outline of research methodology conducted for this research study, research findings, and a discussion of the findings. Finally, it will conclude with a review of next steps and implications for UNC.

LITERATURE REVIEW

This section notes levels of perception researchers have about open access, attempting to distinguish these levels by rank and discipline. It then connects open access perception with researcher publication needs both scientifically and professionally (i.e. in regards to achieving tenure). It will compare the capabilities and limitations of qualitative and quantitative research in these same subjects, and reason why this research study is most appropriate as a bibliometric analysis. Finally, it will look at studies on researcher collaboration, particularly in regards to the identifiable collaboration of "coauthorship."

I found these sources, primarily, through usage of cited reference search of key articles--a search capability that is excellent in index database Web of Science. Initial key articles were discovered using the library's Serial Solutions discovery tool that can search across the various aggregated databases in the UNC catalog. Query terms for the search in research questions 1 and 2 included terms such as "(SUBJ)faculty publishing AND (SUBJ)tenure AND 'open access.'" Query terms in the search in research question 3 included terms such as "(collaboration OR coauthorship analysis) AND (SUBJ)faculty publishing." For the research about open access, I preferred to restrict research study publication dates to 2014-2019, as open access became a much more discussed research focus during this time and the publication methods and models are not as comparable prior to 2014. For research about coauthorship, the publication limit was less important because valuable studies have been occurring since the 1980s.

Setting

The early career phase is where researchers aim to make a name for themselves, conducting research and developing skills to further participate in their fields. A 2015 *Nature* report indicates that the number of tenure-track positions have plateaued and declined, and limited probationary pre-tenure appointments require researchers to take off running (Powell, 2015). Repeatedly, in order to meet the tenure requirements, these researchers behave in more conservative publication choices, preferring the traditional subscription, high impact journals known to have high quality in tenure reviews because the number of publications and citations are frequently counted in tenure considerations. In their research training, early career researchers also face mentor recommendations that advise against straying from these traditional paths. Despite open access increasing citations and views, early career researchers show no more uptake on these venues than their seniors. Nicholas et al (2018)'s study shows an in-depth review of these researchers' needs and beliefs.

The University of North Carolina at Chapel Hill (UNC) sets a good stage for a bibliometric study, for it is a large research intensive university with humanities, social sciences, and STEM disciplines. With much scientific research requiring funding, and many funders, such as the NIH, since 2013 adding a requirement for open access publication of content they had funded, assistant professors are stuck between the traditional choices of what has previously defined quality and these new policies. This study can help to ascertain how UNC faculty are handling these changes. Additionally, UNC is a university that has in 2018-2019 been very active in outreach about open access due to two large subscription bundle licenses up for renegotiation, the big news about

other R1 universities such as University of California engaging in open access, and an active open access-engaged Dean of Libraries, but the effects of these outreach are too soon to tell and prior open access engagement has not before been measured. Therefore, in efforts to measure uptake in this changing environment, understandings about researcher patterns on journal selection become that much more important.

Faculty Engagement of Open Access

Understanding faculty engagement with open access can be dependent upon assessments of awareness of open access, perceptions of open access, and experience publishing in open access. A number of studies reveal a high amount of awareness. In Swan and Brown's study, about two-thirds of their survey participants indicated they knew about open access (Peekhaus, 2015, p. 642). By 2007, 85% of respondents to annual surveys were aware of open access (Xia, 2010). Awareness is a first step, because this can lead to considering open access journals when submitting articles for publication. Once faculty are aware of open access, they can develop perceptions, whether positive or negative.

Positive perception of open access, however, does not directly correlate with publication choice. Direct correlation would be an author who learns about open access and decides to publish open access [or not] because it meets [or does not meet] their needs. But what about that author who believes positively in open access but does not follow through with their actions? Many studies show this disconnect between belief and action (Xia, 2010; UC survey, 2006, as cited in Peekhaus, 2015, p. 642). They identify lack of funding as one reason why authors did not publish open access despite feeling positively about it. In a survey of Canadian engineering scholars, though 83% expressed

support for open access, only 25% had published in an open access journal in the previous two years (as cited in Peekhaus, 2015, p. 642). Rodriguez (2014) provides an excellent overview of research studies in relation to open access perception of scholars. These various studies differ in scope, which makes comparing the results difficult, but they do seem to extrapolate larger tendencies of perceptions, at least between their Humanities, Social Sciences, and STEM fields.

Notably, authors with prior experience publishing in open access are more likely to publish open access again. Peekhaus (2015) notes that findings reveal authors with open access experience have motivations committed to "the principle of free access to research" and perceived the quality to be higher than subscription journals because of the increased readership and citation rates; however, those without open access experience perceive the opposite (Swan & Brown, as cited in Peekhaus, 2015, p.643-44). Park studied across 11 subject areas and found those without open access experience, regardless of rank, felt less confident about publishing in an open access journal. The unfamiliarity of the different publishing model was the cause (Park, as cited in Peekhaus, 2015, p.643).

Studies reveal different findings about rank's influence on a scholar's publication venue choice. According to O'Brien (2011), "Career aging—or rank advancement—theories assert that a scholar's behaviors may change as they meet various milestones in their career" (as cited in Sugimoto et al, 2016, p. 999). Broadly speaking, values on different dissemination formats can change, as may deference in author order so younger authors can take lead author prestige (Sugimoto et al, 2008). In Rodriguez (2014)'s study, age, seniority, and rank were found not to be predictive indicators of a scholar's

awareness, perception, and experience with open access (p. 604). Of all the faculty ranks among the 276 respondents, 74% of the 96 Assistant Professor participants agreed that articles should be free for online access, though the Assistant Professors are less likely to have published in open access journals than Associate Professors (Peekhaus, 2015, p.655). Niles, Schimanski, McKiernan, and Alperin (2019) found that there is a difference in rank between tenure perceptions of journal impact (p. 6, 11). Older faculty (often but not exclusively found to coincide with higher ranks) repeatedly are less likely to adapt to new ideas, so untraditionally recognized forms of publishing such as open access journals may not be highly recommended or viewed by those in the field longer. Notably, older faculty's productivity of cutting-edge research has been found to not decrease after tenure (Sugimoto et al, 2016, p. 998).

Tenure-track Needs and Open Access Publishing

Researcher needs for publication are valuable considerations for publication choice, and can reveal a disparity contrasting with researcher needs for accessing research. Authors prefer relevant journal area coverage, wide circulation of the journal and in the field, and a high impact factor. Older faculty—note, not absolutely related to tenure!—indicate more value in open access qualities of publication than younger faculty do, perhaps because the non-open access incentives are not as relevant to their rank (Ithaka S+R, 2019). There is a difference between what researchers value most and what they think their peers value most. In Niles, Schimanski, McKiernan, and Alperin (2019)'s survey, respondents indicated their primary value for publishing to be journal readership, yet they expected their peers valued prestige and impact factor metrics (p. 7). In Ithaka S+R's 2018 faculty survey, 2/3 of their participants indicated they would be happy if all

subscription journals were flipped open access, yet they still overwhelmingly indicated preference for subscription journals in their publication choices (Ithaka S+R, 2019).

Generally over the past half century, department tenure expectations dismiss or put less weight upon forms of scholarship that are not the peer-reviewed monograph or journal article—this negatively impacts faculty in fields progressing towards public engagement and digital communication from exploring more appropriate dissemination outlets (Working Group on Evaluating Public History Scholarship, 2010). For Public History, the activities scholars are hired to undertake will not fit with tenure criteria. In the Working Group's content analysis of thirty-five diverse universities, the tenure and promotion standards of the time varied from short and informal to long official statements, from static explicit requirements to open-ended options, from unadaptively generic to creatively committed to the conditions of the field. Tenure policies have been expressed by survey participants as unnecessarily constraining their publishing choices. According to Schonfeld and Housewright, one third of their participants strongly felt this way, despite prioritizing only what would be positively reflected in a tenure review (Schonfeld and Housewright, as cited in Peekhaus, 2015, p. 643).

A pre-tenure public health professor stated the strategic play of dissemination choices. "What I try to do is a good balance of really high impact but also journals where I know it could . . . have better influence on individuals" (Hanneke & Link, 2019, p. 344). A tenured professor having participated in a promotion and tenure personnel committee noted their school had begun to expand what "qualified" as scholarly publication—expressing that dissemination outputs that show reach and societal outcome should also be considered impactful (Hanneke & Link, 2019, p. 344). In 1993, a report *Redefining*

Historical Scholarship called to expand the reward beyond monograph publication to meet the needs of the field, and still by 2010 the issue was ongoing and worked to address this by culling a national committee to evaluate the issue (Working Group on Evaluating Public History Scholarship, 2010).

Faculty have expressed desire to disseminate their research in order to do service to the society. This could mean making publications open and free, or choosing outlets like blogs and newsletters that are public facing and understandable to the layperson. For tenure-track faculty, choosing dissemination outlets for these reasons, in addition to publishing in high-impact journals for tenure consideration, is twice the work (Hanneke & Link, 2019, p. 346; Working Group on Evaluating Public History Scholarship, 2010). Perhaps, traditional citations-based impact metrics can be judged in addition to value-based metrics, such as the Becker Medical Library Model for Assessment of Research Impact, which suggests indicators for community benefit, or HuMetricsHSS, which encourages measurement in relational values such as equity (Hanneke & Link, 2019, pg. 346).

Methodology in Qualitative and Quantitative Studies

A number of studies have explored faculty engagement with open access within and across various disciplines, using surveys and semi-structured interviews. It is against these that my study makes analyses differently, because surveys and interviews measure faculty engagement in a different manner than publication data analysis. Though many quantitative findings have been parsed out in qualitative studies—such as the number and statistical breakdown of participants who published open access--human responses collected from qualitative studies can often be subjective to the participant's

understanding of open access publishing, whereas quantitative publication data will reveal what open access publishing has literally occurred. The quantitative results of previous qualitative studies can still be compared to records-based studies like my own, and will be interesting to see similarities and differences. Additionally, qualitative studies can add richness to my results, as they can reveal understandings about faculty engagement with open access based on whether participants *knew* they published open access and what factors influenced them to publish open access (Warlick & Vaughan, 2007). Though my study will not be able to include contextual reasons for faculty publishing choices, my findings can hopefully add accuracy and meaning in conjunction with qualitative studies that use similar sampling.

Selection of participants in survey and interview methods can have much more variation than in a publication data analysis, which affects accuracy of representation (and of research reproduction). Many of these qualitative studies begin by pulling a list from index databases of scholars who have published open access (Warlick & Vaughan, 2007); determining the population begins very similarly to a bibliometric data analysis, but the number of final participants is severely reduced in qualitative studies based on who agrees to participate—this is a limitation of a human study rather than a bibliometric study. In other interview and survey studies, researchers gather their participants from blanket listserv announcements of which no particular person had been initially identified (Dalton, 2010), or from departmental faculty lists, of which no faculty is excluded (Hanneke & Link, 2019; Pho & Tran, 2016; Richardson et al, 2019). The population can be initially narrowed to fit the needs of the research for both qualitative and bibliometric data analyses. Survey populations have the capacity to be random and wide, since they

can reach an enormous pool of participants via a blanket listserv without the researcher having to reach out to individual people, but with less control due to less specificity; and interview studies often use a small pool of participants to [possibly inaccurately or incompletely] represent their population; whereas bibliometric analysis can systematically identify participants within controlled databases.

Qualitative studies cover a variety of sample sizes and types, similar to bibliometric studies: a single institution (Hanneke & Link, 2019; Pho & Tran, 2016), multi-institutional within the same region (Warlick & Vaughan, 2007), unlimited institutions globally (Qualitative: Dalton, 2010) (Bibliometric: Liu & Li, 2018; Miguel, de Oliveira, & Gracio, 2016; Siler et al., 2018), and all programs that have members in an association (Richardson et al., 2019). In bibliometric studies, a huge amount of data, balanced for representation, can be collected at once, so long as publications have been included in a database, but the utility of collecting a large amount of surveys or especially interviews, and getting representative answers, is much more time consuming and dependent upon willing human participants. The initial requests of eligible participants, and the resulting participants in prior literature is often relatively small: about 13 interview participants (Hanneke & Link, 2019; Warlick & Vaughan, 2007), and 150-350 survey responses (Dalton, 2010; Pho & Tran, 2016). The disciplines included in the research studies also varied: thematically chosen, such as biomedical faculty, which consists of a few separate departments (Qualitative: Hanneke & Link, 2019; Pho & Tran, 2016; Warlick & Vaughan, 2007) (Bibliometric: Liu & Li, 2018); topical, such as all researchers who study educational leadership (Qualitative: Richardson et al, 2019) (Bibliometric: Siler et al, 2018); and specific disciplines, such as LIS authors (Dalton,

2010). The previous literature has intentionally narrowed the research focus onto specific disciplines in order to reveal patterns of behavior and engagement among similar research topics; my study will attempt to show representation for all disciplines in one American university, as a case example of representation at a large research intensive institution.

Coauthorship Impact

Collaboration can influence performance of a publication. It can cause increased likelihood of article acceptance, shortened time to promotion, and citation count (Sugimoto et al, 2016, p. 1000). Sugimoto et al. (2016) found that professors collaborate more than assistant and associate professors—they state, "[A]lthough collaboration patterns may be shaped by early experience, collaboration generally becomes relatively more common over time, even within individuals. . . . the least collaborative years are early in the career" (p. 1008). According to Lariviere, Gingras, Sugimoto, and Tsou (2015), 87% of science papers existed as single-authored in 1900, whereas only 7% existed in 2011, and in the humanities and social sciences in the same time span, this number was 97% down to 38% (p. 1330-1; West et al., 2013, as cited in Leahey, 2016). Leahey (2016) cites various papers that study coauthorship in different disciplines. Federal policies and universities are also pushing increased collaboration—UNC's Office of Research Development explicitly exists to support researchers with many coauthors to apply for grants, a concept that Clark and Llorens (2012) also discuss (as cited in Leahey, 2016).

Conclusion

Considering that open access began in the 1990s, the fact that studies even through 2018 have asked faculty about their level of awareness of open access

immediately implies its lack of traction. As an overview, we have pulled out various different factors faculty have about considering open access and are able to see from surveys and interviews how faculty have been found to perceive open access differently depending on their identity in the university. Surveys and interviews are able to collect measurable data, but still reflect the subjective beliefs of the participants--so it is important to note in what ways qualitative and quantitative methods can affect the overall goals and findings. As a counterpart to these belief-based studies and a sample variation against other bibliometric studies, this bibliometric study stakes actual publication choices of UNC tenure-track faculty. Visualizations can only better inform qualitative studies by clearly showing identifiable patterns to respond to for improving communications about publishing for emerging scholars.

METHODS

This study will use quantitative data analysis to assess the proportion of UNC tenure-track faculty members who are publishing in open access journals, and the proportional quantity of open access publications by discipline. This method of analysis will be conducted through the use of bibliometric data. Before setting up the specific steps of this study, it is first important to understand the extent to which this method can help to inform this study's findings.

Quantitative Analysis

Bibliometric data shows what has been done—what has been published, by whom—and will provide no further meaning and context. Quantitative analysis can thus deliver a summary of productivity, such as the proportion of Assistant Professors with open access publications—rather than the factors behind the degree of productivity. If done responsibly, summaries such as these can make meaningful comparisons between and amidst groups and reveal publication patterns. Any limitations in the data collection will naturally limit the extent of such patterns.

To analyze current completed publication activity on campus, quantitative analysis is a more reliable tool than surveys or interviews, because it can factually describe how much and who has published. Survey and interview methods, as revealed in the literature review, can explain qualitatively the difficulties and motivations of publishing activity, but will be subjective data limited to the understanding, awareness, and biases of each participant. Qualitative analysis is a necessary consideration, in

combination to quantitative analysis, for comprehensive understanding, but will not be a method used in this research paper beyond occasional comparison to previous studies in the Discussion section.

Using visualizations of quantitative analyses enables readers to conceptually visualize the extent of these relationships, and, especially if the visualization is interactive, to engage within the related information.

Data Collection

This study will use bibliographic records as its dataset to conduct the quantitative analysis. The dataset included various values of metadata elements within each publication's record--most importantly, UNC Author Name, Title, Publication Year, Journal Name, and whether the article was open access or not. These elements appeared as column names, and each row was a distinct article title.

The Population

To begin data collection, I first needed a list of authors to refer to. I collected this list from UNC's System Database, which is publically available available on the *News & Observer*'s website (https://www.newsobserver.com/news/databases/public-salaries/article11863496.html), retrieved in November 2019. The list, provided by UNC General Administration, contains name, department, title, salary, and hire date for all professors, administrators, and staff employed at UNC during the 2018-2019 academic year. To spot check the quality of this list, I found an equivalent set of current faculty listed in the UNC Directory, in the Advanced Search after selecting the "faculty" field in the dropdown: (https://directory.unc.edu/dirSearch/view.htm); however, this list would not have been a useful standalone faculty source for quantitative analysis given that no

information is provided besides name, email, and phone. As another method for spot checking the employee list, I visited a few individual UNC departmental websites and compared the faculty listed. This source, too, would not have been an efficient standalone faculty source for quantitative analysis since I would have had to manually create my own list. Because the *News & Observer* employee information is public knowledge and open to anyone, I have included the employee population dataset as an open data supplemental file to this paper.

Data Source

I also needed to select a data source from which to collect the employees' article publication information. In ILS research, publication record data are typically collected from bibliographic, or index, databases--I chose to use Scopus. Disciplinary coverage of publications varies depending upon the database source used; Scopus contains a much larger collection of medical and biological publication records than humanities records. More depth on the particular distinctions between various index databases can be found in various studies that compare index database coverage (Falagas et al., 2008; Harzing, 2016; Mongeon & Paul-Hus, 2016; Powell, 2017; Waltman, 2018). Throughout my Results section, I sought to note various findings that may have been affected by Scopus's disciplinary coverage.

To avoid database coverage selection limitations, some previous studies have collected their data by scraping departments' websites for faculty CVs and then creating datasets using the publications listed in the CVs (Sugimoto, 2011; Sugimoto, 2016)--however, limitations for this method of data collection include the lack of faculty CVs posted on departmental websites, and whether the CVs have been updated. Because of

Scopus's known quantity and explained criteria for its journal coverage (Scopus, 2020), the ease of using a single source for data collection, the standardized bibliographic metadata across journals and publishers, and the user friendly API service (as discussed in the below API subheader), Scopus was my preferred data source.

As of October 2019, Scopus includes publications from approximately 40,000 journals, over 5,500 journal titles of which are open access journals, which have been authorized as open access journals if they appear in both or either the Directory of Open Access Journals (DOAJ, at https://doaj.org/) or the Directory of Open Access Scholarly Resources (ROAD, at http://road.issn.org/). In comparison, in October 2019, DOAJ listed a total of 13,865 open access journals, so using Scopus for data collection will have incomplete publication coverage. Scopus also captures grant funding metadata if the funding body is included in the FundRef ontology (http://www.crossref.org/fundref/)--so any analysis about funding should note that it may refer to incomplete data. This was another reason I have chosen in my paper to not focus on funding relevance to open access.

Data Cleaning Prior to Publication Record Data Collection

Since my research questions refer to tenure-track faculty, I first removed any faculty members in the employee population dataset who did not have a tenure-track title. As described in the Introduction, UNC's Office of the Executive Chancellor and Provost policies define tenured faculty at the Associate Professor rank, tenure-track faculty as "instructors" and "assistant professors," and non-tenure track faculty as "teaching professor" or "clinical assistant professor." Therefore, the only faculty left in my UNC

faculty dataset were Assistant Professors and Instructors, not including any who had "clinical" or "teaching" within these titles.

In preparation for collecting publication record data, I needed to ensure I correctly matched the faculty in my employee population dataset to author names within the Scopus database. After all, databases often include authors who share the same name with another author. Scopus assigns authors a Scopus ID, which helps to allocate an article to a specific person. All authors within Scopus will therefore have a Scopus ID. I collected a list of names and Scopus IDs of the 33,528 UNC affiliated authors (through a Scopus affiliation search of "The University of North Carolina at Chapel Hill") and ran code to join the datasets together so that faculty names from my employee population dataset were added on the same row with the matching author name from the Scopus list. Faculty from the employee population dataset who did not have a Scopus ID were not included in the sample. UNC-affiliated authors from the Scopus list who were not included in the employee population dataset were also not included in the sample.

A limitation with the Scopus ID is that it is imperfectly assigned. Scopus automatically assigns author IDs upon any new author's first article in Scopus, but authors sometimes end up with more than one Scopus ID. Not including all of the IDs for a faculty member results in an incomplete publication record collection for that author. Also, it does not always recognize that an author isn't someone else, so in some cases, the matching author name was not actually from UNC. After hiding faculty names from the employee population dataset who did not have a Scopus ID--i.e., were not in Scopus--I reviewed the list to see if the first, middle, and last names were identical or if any faculty were listed on more than one row (as joining datasets can cause spurious tuples). For any

rows that had matching issues, I searched for the names in Scopus and replaced the Scopus author with the correct Scopus ID or removed the faculty member if the correct person did not have a Scopus ID. For faculty that had more than one row, I ensured the names were correct, and kept the extra rows so that multiple Scopus IDs could be searched in my API call.

I assigned a disciplinary category for each UNC employee still in my sample. The disciplinary categories included Humanities and Social Sciences (HSS), Medicine (MED), and Natural Science (NAT). These categories were assigned according to how UNC School of Arts and Sciences broke down departments into categories.[3] Many of the MED departments come from a variety of distinct UNC Schools, such as the School of Pharmacy. For a few departments separate from the UNC School of Arts and Sciences and clearly not in a MED category, I placed them within HSS, such as the School of Education, Law, Information and Library Science, and Social Work.

Publication Records via an API Call

I made an API call to retrieve Scopus publication records, because it was a much faster and scalable method than manually exporting 2,000 publication records at a time. It also would make the study's data collection more accessible to reproduce. As Wildemuth points out, studies collecting data from nontransparent "black box" sources can often cause the study to not be as reproducible, since data selection and access may not be known (Wildemuth, 2017, pg 344). To avoid this, I used an API call that would pull all the data I needed. Not all databases offer API services, but Scopus does.

I secured my authorization key from Scopus Developers (https://dev.elsevier.com/), and referred to documentation about how to make the call

using the ‘rscopus’ library package in R (Muschelli, 2019). By utilizing the author_df function with the list of Author IDs as a parameter, the API call was able to retrieve publication title information from each author. The data was downloaded from Scopus API between January-February 2020 via http://api.elsevier.com and http://www.scopus.com.

The publication record dataset retrieved from my API call included data for title, title ID, corresponding author, journal name, journal ISSN, journal eIssn, volume, issue, page range, publication date, cited by count, publication type, author count, UNC author’s first name, UNC author’s last name, UNC author’s Scopus author ID, article DOI, abstract, PubMed ID, author keywords, open access boolean, and funding source. The API call can be found in my R language programming file that is supplied openly as a supplemental file; it will not include my authorization key so as not to compromise Scopus terms of service.

Data Cleaning After Publication Record Data Collection

Cleaning collected data before making analyses reduces the possibility of making inaccurate or incomplete research findings. After the API call, I replaced any duplicate Scopus IDs with only one ID per faculty member in the UNC faculty dataset and in the publication record dataset, so that in the data analysis, a faculty member could be identified uniquely. I also reviewed the journals included in the dataset to check that journal names were listed uniquely. Since some journal metadata did not include the journal’s ISSN or eISSN, I created a new column to help ensure the unique identity of journals--so that in analysis I would not be using a journal variable that had null values.

Using Tableau, I joined the identified UNC faculty dataset and the publication record dataset. This merged dataset included a separate published article within each row, with publication record and author information in the columns. I removed articles that had null employee values. I removed articles that were published in 2020, since publication counts would be incomplete for the year and since article authors in 2020 may not still be working at UNC. I also removed articles that had been published by the faculty member prior to their hire date at UNC.

Technicalities of data cleaning also include ascertaining that the values are consistently standardized grammatically: Any time an ampersand (&) was used in the publication metadata, the retrieved dataset exploded that row, pushing the following cells in the row an extra cell to the right. Checking that the cells were in their proper positions would be critical to the analyses to even begin to work.

Data Analysis

I used Tableau and, briefly, Excel to analyze my bibliometric data. I created exploratory visualizations in Tableau to compare and explore various data variables, in particular to look at variables apply to open access. For disciplinary analyses, I made efforts to compare articles within their distinct categories, so as to avoid meaninglessly comparing disciplines that function and thus publish differently and to show more representative results. To counteract uneven faculty representation, I often analyzed rates of publishing rather than total publication counts. Findings can reveal patterns of Assistant Professor article publication choices, overall and distinguished by discipline, as they relate to open access.

Ethics Statement

This study has no foreseeable ethical concerns: my UNC subscription to Scopus granted me access to use its API service and aggregate the bibliometric data for research study and noncommercial use. Faculty information in the Employee population dataset is publicly available, since UNC is a public university and thus required to post such information. I have made efforts to promote reproducible research: By following the methods I describe, researchers should have little trouble creating a similar Scopus dataset to my own.

RESULTS

The Sample

Among the 978 tenure-track faculty (Assistant Professors and Instructors) at University of North Carolina at Chapel Hill, the data source I used, Scopus, had publication records listed for 300 of them. I thus have chosen to use these 300 faculty as the sample.

The disciplinary categories (HSS, Natural Sciences, and Medicine) represented within the sample were most weighted towards Medicine: 200 faculty were in medicine disciplines, 64 were in Natural Science disciplines, and 12 were in Humanities and Social Science disciplines. The most frequent departments in the sample included School of Pharmacy's Division of Pharmacotherapy and Experimental Therapeutics (11 faculty) and Chemistry (10 faculty). The most frequent departments that were entirely not in the sample included School of Education (9 faculty) and Music (7 faculty); the most frequent departments with faculty who were not in the sample despite others in the department being in the sample included Kenan-Flagler Business School (39 faculty) and General Anesthesia (30 faculty). As only one department was listed for any faculty in the employee dataset, any existing cross-departmental affiliations have not been counted in analysis.

Faculty members in the sample published in 5.9% of the total journals represented in Scopus. The journals used by the most faculty in the sample were *PLoS One* (70 faculty), *Scientific Reports* (31 faculty), and *Nature Communications* (25 faculty). The

journals used for the most articles by faculty in the sample were *Lecture Notes in Computer Science* (99 articles), *PLoS One* (92 articles), and *Journal of Athletic Training* (55 articles).

Proportion of Open Access Publishing by Faculty

239 out of the 300 faculty (80% of the sample) had published at least one open access article. Among those who had published at least one open access article, the average number of open access articles published by faculty members is 7.3. In comparison, all faculty members in the sample published on average 11 non-open access articles. On average, faculty members published 40% of their total number of article publications as open access articles. This rate thus shows the extent of open access adoption by UNC faculty. Figure 1 shows the top 10 faculty members who have the highest proportion of open access articles (so long as they published 5 or more articles). Figure 2 shows the top 10 faculty members who have published the most total articles as well as what proportion of these publications were open access. Note that the top 10 faculty members in these figures do not overlap, which means that the top productive authors still publish most of their articles non-open access. However, as many of the authors in Figure 2 are publishing about 20-50% of their works open access, and the authors in Figure 1 repeatedly publish their work open access, their adoption of open access would seem intentional.

Figure 1: Employees with the highest rate of open access publishing, compared to their total article count

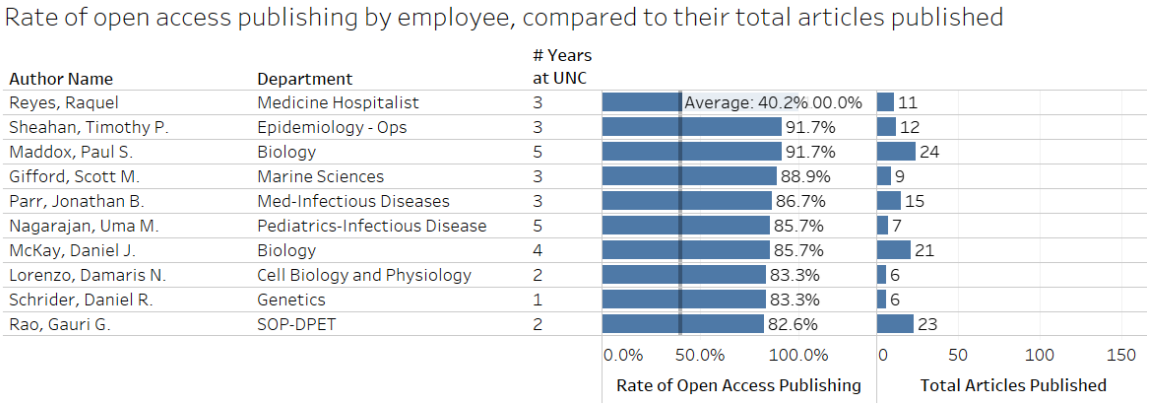
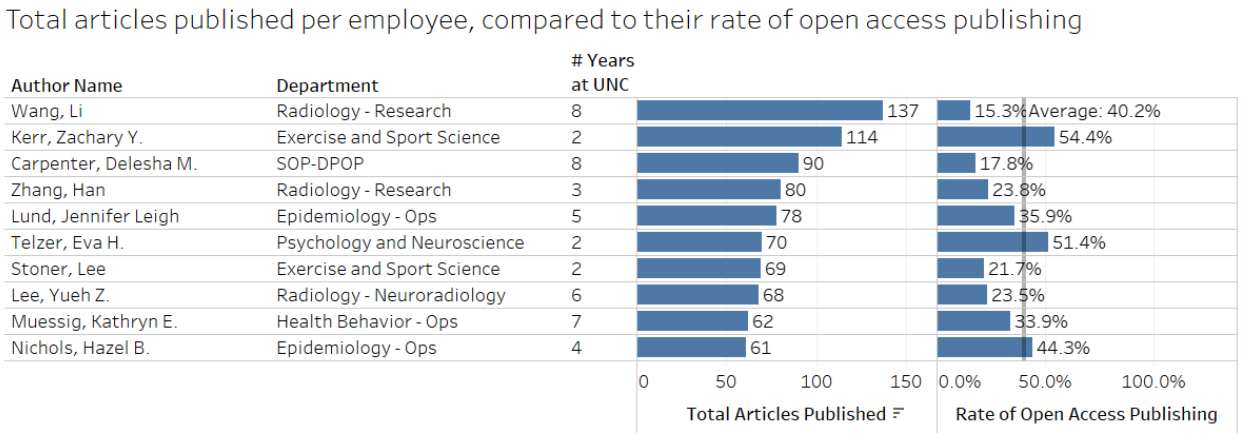


Figure 2: Employees with the most published articles, compared to their rate of open access publishing



Looking at the disciplinary proportions among those faculty in the sample, of the 171 Medicine faculty publishing open access (85% of this discipline's sample), on average they published 7 open access publications compared to an average of 16 total publications. Of the 60 Natural Science faculty publishing open access (93% of this discipline's sample), on average they published 8 open access articles compared to an average of 18.5 total publications. Of the 12 Humanities and Social Science faculty

publishing open access (100% of this discipline's sample), on average they published 4 open access publications compared to an average of 16 total publications

Faculty members published with a rate of 11 coauthors per article for open access articles and 9 coauthors per article for non-open access articles. The coauthor rates for both open access and non-open access remain fairly steady regardless of the publication year of the article.

Proportion of Open Access Publishing by Journal

Collectively, the faculty published their open access articles the most in *PLoS One* (89 articles). Figure 3 shows a list of the journals that faculty used most frequently to publish open access. In comparison, Figure 4 shows the journals from 2015-2017 that had the most open access publications within all of Scopus (not just this sample). Note that three of the journals are in both plots. Though *PLoS One* in particular publishes only open access articles, not all journals exclusively publish open access articles, so these lists may include "hybrid journals" that faculty choose often to publish their articles as open access.

Figure 3: Journals with the most open access articles by faculty

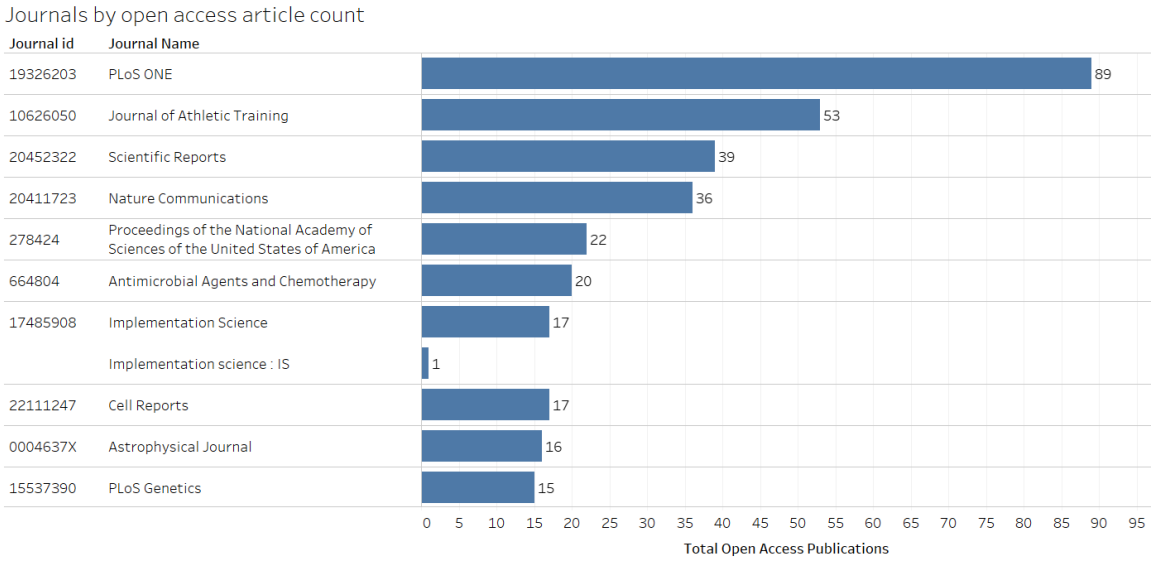


Figure 4: Journals in Scopus overall with the most open access articles in 2015-2017 (Scopus Source List, last updated October 2019. Accessed April 29, 2020. Retrieved from https://www.elsevier.com/solutions/scopus/how-scopus-works/content)

Publication Name	**# Open Access Articles**
PLoS ONE	73,990
Scientific Reports	57,406
RSC Advances	32,913
Oncotarget	19,230
Nature Communications	11,691
Optics Express	9,305
Medicine (United States)	9,198
The BMJ	9,196
IFAC-PapersOnLine	8,187
BioMed Research International	7,395

Proportion of Open Access Publishing by Discipline

Nearly equally, three of the top five departments overall publishing open access articles are from the natural sciences. This likely can be identified due to the few prolific natural science faculty members shown in Figure 2.

Within the sample, 70% of faculty publishing open access come from a medicine discipline. Therefore, due to the sample's high number of faculty in the medicine category, I have chosen to calculate the data by percent of total within individual categories rather than across all categories in an effort to make the relationships more fairly balanced. Figure 5 (Humanities and Social Science), Figure 6 (Medicine), and Figure 7 (Natural Science) each show the top 5 departments publishing open access articles. The online interactive visualization shows more details for each discipline,

including the number of titles and the number of journals published in, and can be drilled down by open access and non-open access publications (go to Figure 5,6,7 tabs at https://public.tableau.com/profile/lynnee.argabright#!/vizhome/OpenAccessPatterns_LArgabright/Figure1).

Figure 5: Humanities and Social Science departments publishing open access articles

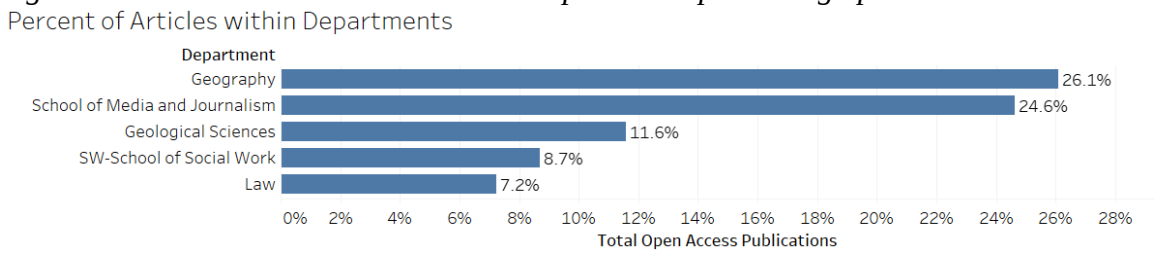


Figure 6: Medicine departments publishing open access articles

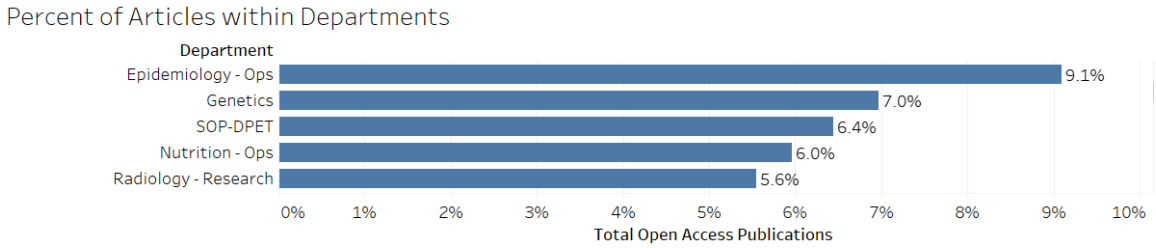
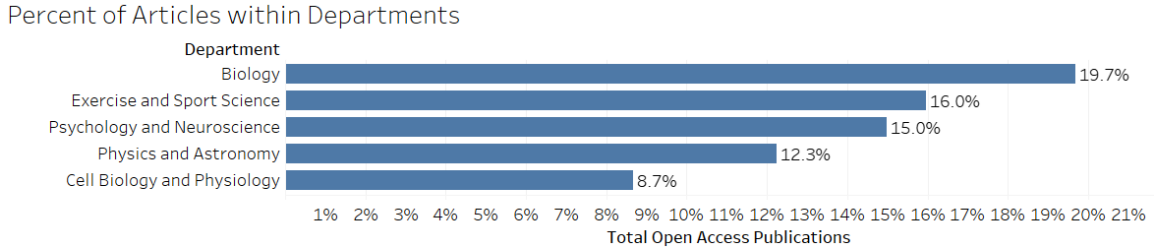
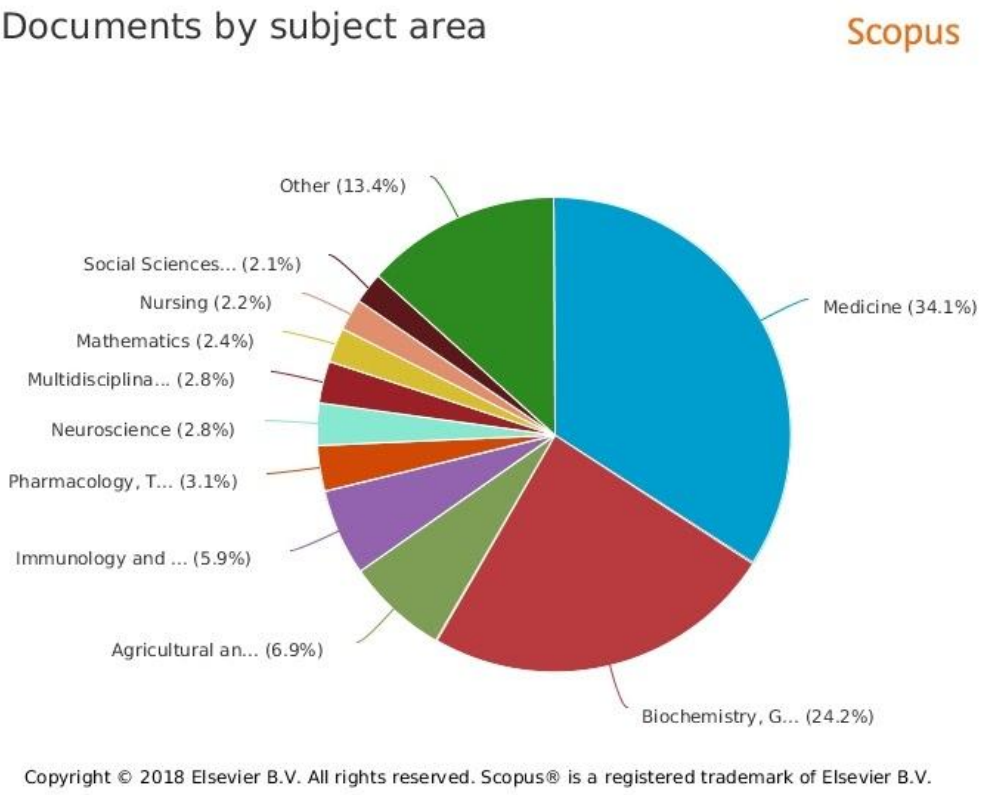


Figure 7: Natural Science departments publishing open access articles



Medicine subject area journals that publish open access articles represent an overwhelming majority within Scopus. Figure 8 shows the proportion of open access subject areas in Scopus overall. Note that the proportions of subject areas represented in Scopus in Figure 8 are similar to the proportions found within the sample dataset.

Figure 8: Open access journals in Scopus by subject area (Credit: Elsevier)



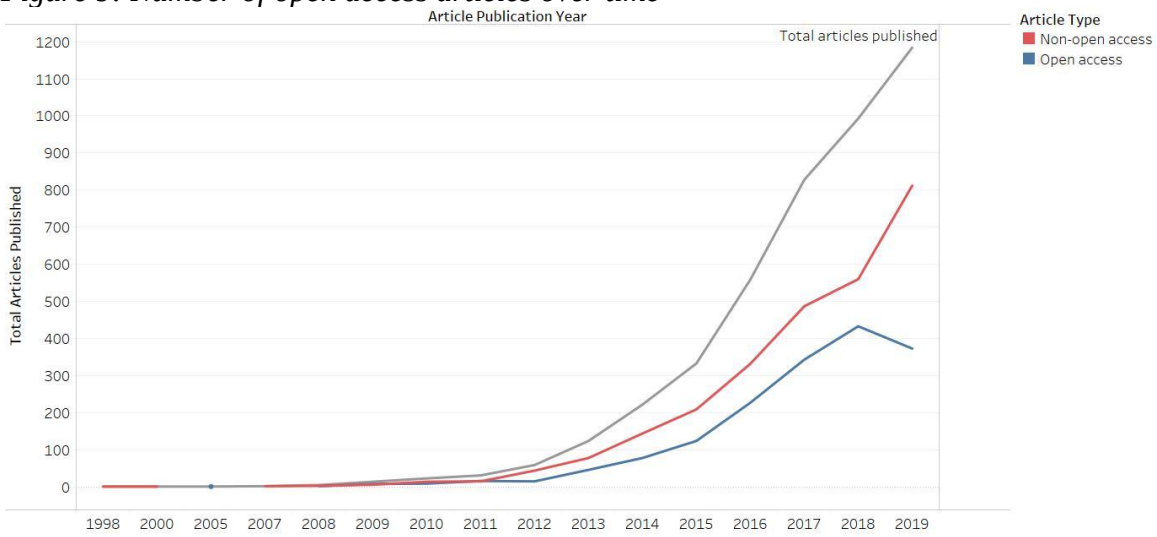
Open Access Publishing Citation Impact

Within the sample's publications, the open access articles received on average more citations than the non-open access articles. Though open access is not the only variable impacting citation count, it is worth noting. The average citation count for open access articles in the sample was 14.3 citations per article, whereas the average for non-open access articles was 12.6 citations per article. However, these average citation counts could be affected by a scant few articles that have much higher citations than the rest, as the majority of articles have 150 citations or fewer and the most cited articles had 819 and 5,499 citations. If those top articles are removed from the analysis, the average citation count becomes 13.7 citations per open access article and 10.6 citations per non-open access article.

Open Access Publishing Over Time

In the sample, tenure-track faculty began to publish more articles starting in 2013. This is partly because there were only 47 out of 276 assistant professors and instructors (who are still assistant professors and instructors as of academic year 2018 when they were added into the dataset) who were employed at UNC prior to 2013, so the number of articles would naturally increase if six times as many authors were publishing. Since then, the number of open access and non-open access articles grew at a similar pace, with about 100 fewer open access articles each year between 2015-2018. From 2018-2019, the number of open access articles decreased by 60 articles, whereas the non-open access articles increased by 252 articles. Possible rationales causing this gap are in the Discussion section. Also, between 2018-2019, 34 more authors published non-open access articles and 6 fewer authors published open-access articles. This does not necessarily mean that the 60 fewer open access articles are directly caused by those 6 fewer authors. Figure 9 shows the progression of faculty article publications over time--showing total articles in the top line (grey), compared to the split between non-open access articles in the middle line (red) and open access articles in the bottom line (blue). The online interactive visualization shows a breakdown of articles that were published in these distinct years (go to Figure 9 tabs at https://public.tableau.com/profile/lynnee.argabright#!/vizhome/OpenAccessPatterns_LArgabright/Figure1).

Figure 9: Number of open access articles over time

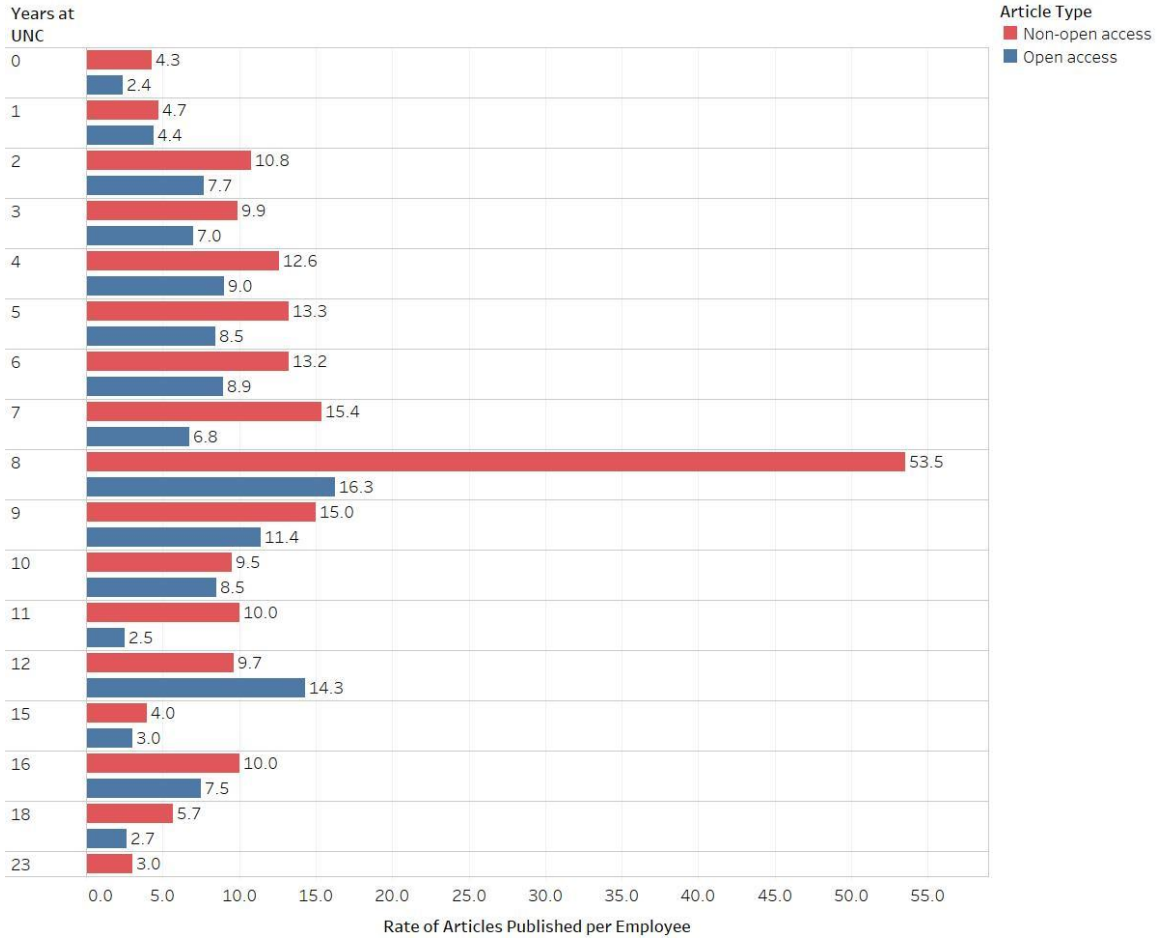


According to the sample data, the number of years an assistant professor or instructor has been at UNC may have an effect on the number of publications they produce. Figure 10 shows the rate of articles (split between open access and non-open access) that faculty members publish depending on how long they have worked at UNC. For example, those who have worked at UNC for 1 year (a 2017 UNC hire date) have published at a rate of 9 articles per person since being hired in 2017, and those working for 2 years (a 2016 UNC hire date) have published at a rate of 19 articles per person since being hired in 2016. To clarify, this figure does not describe *when* the publications occur, but rather shows an assistant professor's total rate of publication history since they were hired as an assistant professor at UNC. There seems to be similar publication rates between those hired in neighboring years: the rates are fairly consistent between year 0-1, 2-3, and 4-6; however, the circumstances for why faculty publish differently across years would require qualitative analysis. The figure follows a sort of bell curve--as the employment duration increases, faculty members' average rate of publication increased up to those in year 7, and then the rate decreased for faculty who have worked longer

than 7 years. It is important to mention that as UNC has few assistant professors and instructors working under this title for more than 6 years, the average rate fluctuated more for these longer employment duration categories. This is particularly relevant for the outlying publication rate of faculty members who have worked at UNC for 8 years--only four people are represented in this category, compared to 42 people represented in the sample who are in their second year at UNC, so these four people's individual productivity caused a high publication rate for the category. That being said, though, faculty members who have worked at UNC longer are likely to have increased their rate of publications over time rather than maintaining a steady publication rate, as the overall number of publications has increased over time (see Figure 9). For details about the publication years of articles by faculty members in these categories, the online interactive visualization shows a breakdown of faculty members' individual articles (go to Figure 10 tabs at https://public.tableau.com/profile/lynnee.argabright#!/vizhome/OpenAccessPatterns_LArgabright/Figure1).

Also, across the categories of employment duration at UNC, faculty members publish lower rates of open access articles compared to non-open access articles. This trend appears to follow the difference in open access and non-open access publication totals in Figure 9.

Figure 10: Open access and non-open access publication rate depending on the number of years at UNC



DISCUSSION

80% of the sample having published at least one open access article may seem high, except that the sample primarily consisted of medicine and natural science discipline faculty. These disciplines usually correlate with receiving grants to do research; in fact, 82% of the faculty in the sample had been able to secure at least one grant. Grants often provide an incentive to make the research open access--either by stipulating in the acceptance of the grant that the funded research outputs must be made open, or by using a portion of the budgeted funding to pay an open access Article Processing Charge. Setting aside the disciplinary imbalance of the sample, among those who had published open access, they on average published notably more than 1 article. This suggests that much of this 80% of open access publishing authors were, firstly, aware of the existence of open access, and secondly, that their choice to publish open access was intentional. As many of the authors in Figure 2 are publishing about 20-50% of their works open access, and the authors in Figure 1 repeatedly publish their work open access, they are likely to continue publishing open access in the future. Qualitative research on faculty awareness and motivations to publish open access can be found in the Literature Review section.

The 40% proportion of open access articles (compared to non-open access articles) within the sample reflect the findings from qualitative studies discussed in the Literature Review, such as in Peekhaus (2015). When looking at Figure 9, the number of open access articles consistently lagged about 100 articles behind those that were non-

open access, despite both drastically increasing over time. Again, in Figure 2 and Figure 10, the rate of open access articles published within total publications per faculty member is repeatedly lower. During these tenure-track faculty's short 6-year window to compile an impressive tenure review portfolio, they may feel limited when choosing their publication type based on what journals will secure them high impact. As discussed in Alperin et al. (2019), tenure portfolio guidelines do not at this time adequately encourage tenure-track faculty to produce research outputs that have public impact (such as open access content). However, as over a quarter of the sample's overall outputs are still open access, and the number of open access articles did climb at a similar rate to non-open access articles through 2018, the open access trend is visibly growing. Qualitative observation would clarify whether this is due to grant stipulations, departmental acceptance of open access, or personal faculty motivations at UNC-Chapel Hill specifically.

The gap in non-open access and open access articles published between 2018 and 2019, in Figure 9, did not follow the steadily increasing article publications of faculty since 2013. There was little indication within the data that other variables were disrupted that could have caused a decline in open access articles. Based on the only slight decrease of authors publishing open access despite 60 fewer articles being published, the reason appears to be not specifically due to those 6 people no longer publishing open access but due to a change in many of the faculty members' decisions to publish as much open access as they had previously. For that year, the rate of open access publications per faculty went down from an average of 40% to 31.7%. Future research is needed to determine whether a change in publication choice decision is due to larger administrative

or political reasons, such as through data source embargoes, journal Article Processing Fee changes, and amount of grants available. Specific policies or disruptions at UNC could have incentivized faculty to choose to publish non-open access articles. Though I investigated news and policies surrounding UNC in 2019, I was not able to determine a cause for this drop. A replication of this study at a later date would be enlightening to see if the drop disappears or is only temporary.

Citation count for this particular dataset was less distinctive than prior research studies have exposed. As Piwowar (2018) found, open access articles typically receive a dramatically higher proportion of citations than non-open access articles. I expect that, had I increased the number of publications in the dataset, I would see the proportion gap rising. I think this because including articles published by the faculty member prior to their hire date at UNC causes the dataset to grow from 1,675 to 2,841 and shows a result of 34 average citations per open access article and 25 average citations per non-open access article. Though such articles do not adequately reflect UNC-affiliated publications, further data collection would likely show more accurately representative results.

The frequency distinction between journals most used by faculty and journals with the most articles was interesting. A journal that has multiple authors publishing in it shows diversity of perspective and voice. If a journal has authors from different departments publishing in it, it can become identified as a journal that publishes interdisciplinary research, or it could be known as a high level journal that will publish all subjects without as much depth as more targeted journals. As the representation of departments in the faculty sample was uneven (70% faculty in the Medicine disciplinary category), it would be hard to tell the true extent of these leading journals'

interdisciplinarity, such as whether they cross disciplinary boundaries. A more representative selection of faculty in the sample would be needed to determine journal interdisciplinarity. Conversely, a journal with the most articles published throughout the sample could mean that it is also interdisciplinary, or it is getting many repeat author or department publications. A few of the faculty members in the sample were particularly prolific in their publications--looking specifically at who is publishing all those articles could identify whether those single authors were the only ones producing all those articles for a top frequency journal or whether there were more possible authors who could be contributing all those articles. This is one reason why having a larger sample size would be helpful.

Collaboration

My brief examination of author count suggests there is more to study here, but to do so would require more data. As the sample dataset showed, there was little difference between open access and non-open access publications in terms of bibliometric authorship. Studies on contributorship declare that authorship does not completely encompass acknowledgement of research collaboration. Unfortunately, published articles do not consistently use a standard method of noting contributors, such as CReDIT, nor have any methods beyond authorship been added as metadata into index databases for bibliometric data collection. Data about more diverse acknowledged roles for faculty members involved in publications could show insights about tenure-track faculty research participation, as Cheruvelil et al. (2014) describes. As the rate of article publishing increased for faculty who have worked at UNC between 0-7 years (see Figure 10), even if the average rate of coauthorship did not change much, the collaboration practices may

change. Interviews with assistant professors could reveal whether the type and amount of research changes throughout a tenure-track career, thereby affecting reasons and methods of collaboration.

CONCLUSION

This quantitative analysis served as a case study revealing the state of open access publishing among tenure-track faculty at UNC-Chapel Hill. Based on the data collected, most of the faculty in the sample have published at least one open access article, but on average have published 7 open access articles, since their hire date at UNC.

Limitations and Future Study

While this study seeks to showcase a general picture of faculty output actions, it did have limitations. Firstly, it did not explain the cause of faculty behavior. Using quantitative analysis means the results can reveal patterns of faculty behavior, but can only describe what occurred. Though the literature review has provided context for qualitative research on faculty publishing behavior, it cannot specifically speak to the behavior of tenure-track faculty at UNC-CH. For this, interviews or surveys from these faculty on campus would be most authoritative to give insight into why they chose specific journals and made decisions to publish open access.

Secondly, as the Results section breaks down descriptive statistics about the sample, this study did unevenly provide more STM representation than Humanities and Social Sciences, which will affect the proportionate outcomes. This is in large part a limitation of the data source--not all 978 assistant professors and instructors at UNC were listed in Scopus, so I could only use the 300 that Scopus did include. Because Scopus has a focus on science research, there would naturally be much fewer Humanities and Social Science faculty listed in the database. Similarly, the publications by the faculty that I

pulled from Scopus were limited to articles within the journals that Scopus has chosen to include in its index. Therefore, even if a faculty member had a Scopus ID, not all of their publications during their employment at UNC may have been able to be collected. I have sought to point out opportunities for more accurate representation throughout the findings, which can be considered upon future studies. In the meantime, to mediate disciplinary representation in analyses about the data collected, I used average rates or percentages rather than total counts and provided breakdowns of the results at a disciplinary categorical level.

Thirdly, my third research question about faculty collaboration could only be answered briefly due to the limitation of using Scopus's API. When collecting the data through the API, I was unable to collect the coauthor names and affiliations, which I could have retrieved had I manually exported the metadata through the Scopus user interface. Further study on who the coauthors are would provide more information about the extent of collaboration in open access publishing. This could show which institutions (and even countries) are collaborating with UNC assistant professors, how much interdisciplinary research UNC assistant professors are doing, and what the authors' positions are. Collaboration is a valuable consideration in terms of academic productivity and innovation, and could suggest the current potential for assistant professors.

Implications

Knowing the state of open access publishing by tenure-track faculty can help universities to know how research is being disseminated in this new era of scholarly communication. Academic librarians, in particular, can use this research to better understand faculty activity and make more targeted outreach about how to increase open

access scholarship, by connecting with faculty "open access ambassadors" to help promote awareness about how and why to publish open access, and by reaching out to departments with impactful data about its publishing trends. The study can inform efforts to provide scholarly communication services for tenure-track faculty and emerging scholars. Though the data is composed of publications by UNC-Chapel Hill researchers, the methods could be transferable to learning about the state of open access publishing at other R1 institutions. Recognizing faculty publishing trends at large through this research should only be applied in regards to other tenure-track faculty at similarly sized R1 institutions.

Though not included in this iteration's data collection, a future bibliometric study about UNC could be particularly imminent; within the 2019-2020 academic year alone, University of North Carolina at Chapel Hill has experienced two major changes to academic publishing that are likely to increase open access publishing in the future.

First, the University Libraries cancelled both Elsevier (effective May 2020) and Sage (effective January 2020) "Big Deals," which both point to different outcomes. The cancelled access to future articles published in a substantial portion of Elsevier journals is likely to make faculty members personally realize subscription content's (that is, non-open access) negative impact for research (Blouin & Westbrooks, 2020). The libraries, led by Elaine Westbrooks, have done much outreach to UNC departments and administrators about the matter, so that faculty can become aware about current problems in scholarly communication and can know about some alternative forms of access and publishing (such as open access). Though Sage research journal bundles have not been cancelled like those from Elsevier, they include a transformative open access fund.

Because of this, the libraries will have funding available for early-career researchers to publish open access articles in Sage journals (UNC University Libraries, 2019). It will be up to the success of this funding's outreach on campus to determine how effectively this promotes open access publishing.

A second major change to academic publishing at UNC-CH has been the coronavirus epidemic. The need for accessing epidemic-related research has not only created incentive for faculty to make their research outputs (including data and data visualizations) open and distributable (Hudson, 2020), but has caused publishers to open up relevant journal content temporarily or permanently (Kramer, 2020). I expect at least a large uptick in open access publications in 2020 as a result of faculty seeing the critical value of open access epidemic research.

At large, the scholarly community has made progress towards open access scholarship. As shown in this study, UNC is contributing to this. Every bit helps to make a public impact.

NOTES

1. Ranks defined by UNC-CH's Office of the Executive Chancellor and Provost. Retrieved from https://academicpersonnel.unc.edu/faculty-policies-procedures-guidelines/faculty-appointments/

2. For example, UNC-CH's Tenure and Promotion evaluation policies can be found at https://academicpersonnel.unc.edu/faculty-policies-by-school/. The School of Information and Library Science (SILS) does not require a specific amount or balance of publications to confer a promotion, though it attempts to outline the variety of possible scholarly dissemination types such as refereed books and even grant awards. The Kenan Flagler Business School (KFBS), however, specifically lists the accepted tiers of top core disciplinary journals and strongly encourages their faculty to have published within these to be considered.

3. UNC School of Arts and Sciences departmental breakdown by category: https://college.unc.edu/news-and-features/departments-curricula-centers-institutes/#pane-0-1

BIBLIOGRAPHY

Agarwal, A., Durairajanayagam, D., Tatagari, S., Esteves, S. C., Harlev, A., Henkel, R., … Bashiri, A. (2016). Bibliometrics: Tracking research impact by selecting the appropriate metrics. *Asian Journal of Andrology*, *18*(2), 296–309. Retrieved at https://doi.org/10.4103/1008-682X.171582

Alperin, J. P., Nieves, C. M., Schimanski, L. A., Fischman, G. E., Niles, M. T., McKiernan, E. C. (2019). Meta-Research: How significant are the public dimensions of faculty work in review, promotion and tenure documents? *eLife*. Retrieved at https://doi.org/10.7554/eLife.42254

Beasley, G. (2016). Article processing charges: A new route to open access? *Information Services & Use, 36*(3/4), 163–170. Retrieved at https://doi.org/10.3233/ISU-160815

Blouin, B., & Westbrooks, E. (2020, April 9). Upcoming Elsevier cancellations [News Report]. UNC University Libraries. Retrieved at https://library.unc.edu/2020/04/upcoming-elsevier-cancellations/

Cheruvelil, K. S., Soranno, P. A., Weathers, K. C., Hanson, P. C., Goring, S. J., Filstrup, C. T., & Read, E. K. (2014). Creating and maintaining high-performing collaborative research teams: The importance of diversity and interpersonal skills. *Frontiers in Ecology and the Environment, 12*(1), 31-38. doi:10.1890/130001

Dalton, M. (2013). A dissemination divide? The factors that influence the journal selection decision of Library & Information Studies (LIS) researchers and

practitioners. *Library and Information Research, 37*(115), 33–57. Retrieved at http://hdl.handle.net/10197/4956

Falagas, M. E., Pitsouni, E. I., Malietzis, G. A., & Pappas, G. (2008). Comparison of PubMed, Scopus, Web of Science, and Google Scholar: Strengths and weaknesses. *FASEB Journal, 22*(2), 338-342. doi:10.1096/fj.07-9492LSF

Hanneke, R., & Link, J. M. (2019). The complex nature of research dissemination practices among public health faculty researchers. *Journal of the Medical Library Association : JMLA, 107*(3), 341–351. Retrieved at https://doi.org/10.5195/jmla.2019.524

Harzing, A., & Alakangas, S. (2016). Google Scholar, Scopus and the Web of Science: A longitudinal and cross-disciplinary comparison. *Scientometrics, 106*(2), 787-804. doi:10.1007/s11192-015-1798-9

Hudson, S. (2020, April 13). Spreading coronavirus research. *The Well*. Retrieved at https://thewell.unc.edu/2020/04/13/spreading-coronavirus-research/

Ithaka S+R US Faculty Survey 2018. (2019). Retrieved September 22, 2019, from Ithaka S+R website: https://sr.ithaka.org/publications/2018-us-faculty-survey/

Kenan-Flagler Business School (2017). GUIDELINES FOR REAPPOINTMENT, TENURE AND PROMOTION. Retrieved at https://academicpersonnel.unc.edu/files/2018/02/2017-KFBS-Guidelines-for-Reappointment-Tenure-and-Promotion.pdf

Kramer, B. (2020, April 13). Licences and permanence of COVD19 collections in PubMed Central (PMC). Retrieved at https://github.com/bmkramer/COVID19-PMC-collections

Larivière, V., Gingras, Y., Sugimoto, C. R., & Tsou, A. (2015). Team size matters: Collaboration and scientific impact since 1900: On the Relationship Between Collaboration and Scientific Impact Since 1900. *Journal of the Association for Information Science and Technology*, *66*(7), 1323–1332. Retrieved at https://doi.org/10.1002/asi.23266

Leahey, E. (2016). From Sole Investigator to Team Scientist: Trends in the Practice and Study of Research Collaboration. *Annual Review of Sociology*, *42*(1), 81–100. Retrieved at https://doi.org/10.1146/annurev-soc-081715-074219

Liu, W., & Li, Y. (2018). Open access publications in sciences and social sciences: A comparative analysis. *Learned Publishing, 31*(2), 107-119. doi:10.1002/leap.1114

McKiernan, E. C. (2017). Imagining the "open" university: Sharing scholarship to improve research and education. *PLoS Biology; San Francisco*, *15*(10). Retrieved at http://dx.doi.org/10.1371/journal.pbio.1002614

Miguel, S., Tannuri de Oliveira, E., & Cabrini Grácio, M. (2016). Scientific production on open access: A worldwide bibliometric analysis in the academic and scientific context. *Publications*, *4*(1), 1. doi:10.3390/publications4010001

Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of web of science and scopus: A comparative analysis. *Scientometrics, 106*(1), 213-228. doi:10.1007/s11192-015-1765-5

Muschelli, J. (2019, Sept 17). Package 'rscopus': Scopus Database 'API' Interface. Version 0.6.6. Retrieved from https://cran.r-project.org/web/packages/rscopus/rscopus.pdf

Nicholas, D., Herman, E., Xu, J., Boukacem-Zeghmouri, C., Abdullah, A., Watkinson, A., . . . Rodríguez-Bravo, B. (2018). Early career researchers' quest for reputation in the digital age. *Journal of Scholarly Publishing, 49*(4), 375-396. doi:10.3138/jsp.49.4.01

Niles, M. T., Schimanski, L. A., McKiernan, E. C., & Alperin, J. P. (2019). Why we publish where we do: Faculty publishing values and their relationship to review, promotion and tenure expectations. *bioRxiv*. Retrieved at https://dx.doi.org/10.1191/706622

Peekhaus, W., & Proferes, N. (2015). How library and information science faculty perceive and engage with open access. *Journal of Information Science*, *41*(5), 640–661. Retrieved at https://doi.org/10.1177/0165551515587855

Pho, P. D., & Tran, T. M. P. (2016). Obstacles to Scholarly Publishing in the Social Sciences and Humanities: A Case Study of Vietnamese Scholars. *Publications; Basel*, *4*(3), 19. Retrieved at http://dx.doi.org./10.3390/publications4030019

Piwowar, H., Priem, J., Lariviere, V., Alperin, J. P., Matthias, L., Norlander, B., Farley, A., West, J., & Haustein, S. (2018). The state of OA: A large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ, 6*:e4375. Retrieved at https://doi.org/1.7717/peerj.4375 .

Powell, K. (2015). The future of the postdoc. *Nature, 520*(7546), 144–47. Retrieved at https://www.nature.com/news/the-future-of-the-postdoc-1.17253

Powell, K. R., & Peterson, S. R. (2017). Coverage and quality: A comparison of Web of Science and Scopus databases for reporting faculty nursing publication metrics. *Nursing Outlook, 65*(5), 572-578. doi:10.1016/j.outlook.2017.03.004

Richardson, J. W., McLeod, S., & Hurst, T. (2019). Perceptions of Educational Leadership Faculty Regarding Open Access Publishing. *International Journal of Education Policy and Leadership, 15*(5). Retrieved at https://doi.org/10.22230/ijepl.2019v15n5a817

Rodriguez, J. E. (2014). Awareness and Attitudes about Open Access Publishing: A Glance at Generational Differences. *The Journal of Academic Librarianship, 40*(6), 604–610. Retrieved at https://doi.org/10.1016/j.acalib.2014.07.013

Scopus. (2020). Content Coverage Guide [Report]. *Scopus*. Retrieved from https://www.elsevier.com/solutions/scopus/how-scopus-works/content

Siler, K., Haustein, S., Smith, E., Larivière, V., & Alperin, J. P. (2018). Authorial and institutional stratification in open access publishing: The case of global health research. *Peerj, 2018*(2), e4269. doi:10.7717/peerj.4269

Suber, P. (2012). Chapter 2: Motivation. Open Access. Boston, MA: MIT Press. Retrieved at https://openaccesseks.mitpress.mit.edu/

Sugimoto, C. R., Russell, T. G., Meho, L. I., & Marchionini, G. (2008). MPACT and citation impact: Two sides of the same scholarly coin? *Library & Information Science Research, 30*(4), 273–281. Retrieved at https://doi.org/10.1016/j.lisr.2008.04.005

Sugimoto, C. R. (2011). Collaboration in information and library science doctoral education. *Library & Information Science Research, 33*(1), 3–11. Retrieved at https://doi.org/10.1016/j.lisr.2010.05.003

Sugimoto, C. R., Sugimoto, T. J., Tsou, A., Milojević, S., & Larivière, V. (2016). Age stratification and cohort effects in scholarly communication: A study of social

sciences. *Scientometrics*, *109*(2), 997–1016. Retrieved at https://doi.org/10.1007/s11192-016-2087-y

Tennant, J.P., Waldner, F., Jacques, D.C., Masuzzo, P., Collister, L.B., & Hartgerink, C.H.J. (2016). The academic, economic and societal impacts of Open Access: an evidence-based review [version 3]. F1000Research, 5:632. DOI:10.12688/f1000research.8460.3

UNC University Libraries. (2019, Oct 22). Library to debut open access pilot with Sage Publishing [News report]. Retrieved from https://library.unc.edu/2019/10/sage-pilot/

University of California. (2019, August 29). *UC Open Access Tipping Point Public Forum* [Video file]. Retrieved at https://youtu.be/YR34-fzZtUM

Waltman, L. (2018, Oct 3). Web of Science, Scopus, Dimensions, and beyond: The evolving landscape of bibliometric data sources [PowerPoint]. Presentation at the ROI-AV conference: Visuals and Analytics that Matter (Copenhagen, Denmark). Retrieved at https://www.slideshare.net/LudoWaltman/web-of-science-scopus-dimensions-and-beyond-the-evolving-landscape-of-bibliometric-data-sources

Warlick, S. E., & Vaughan, K. T. L. (2007). Factors influencing publication choice: Why faculty choose open access. *Biomedical Digital Libraries*, *4*(1), 1–12. Retrieved at https://doi.org/10.1186/1742-5581-4-1

Wildemuth, B.M. (2017). Chapter 34: Social Network Analysis. *Applications of Social Research Methods to Questions in Information and Library Science.* Second Edition. Santa Barbara, CA: Libraries Unlimited.

Working Group on Evaluating Public History Scholarship (2010). *Tenure, Promotion, and the Publicly Engaged Academic Historian* [White Paper]. Retrieved at https://ncph.org/wp-content/uploads/2010/06/Engaged-Historian-White-Paper-FINAL.pdf

Xia, J. (2010). A longitudinal study of scholars attitudes and behaviors toward open-access journal publishing. *Journal of the American Society for Information Science and Technology*, *61*(3), 615–624. Retrieved at https://doi.org/10.1002/asi.21283

SUPPLEMENTAL FILES & ONLINE VISUALIZATION

Supplemental files posted in the Carolina Digital Repository include:

- ReadMe file
- R language programming file (including API call)
- Employee population dataset

Online interactive visualizations are available for Figures 1, 2, 3, 5, 6, 7, 9, and 10. They show the figures in a larger view, and include additional relevant information when clicked or moused over. These can be found at https://public.tableau.com/profile/lynnee.argabright#!/vizhome/OpenAccessPatterns_LArgabright/Figure1